Statistics & Probability Letters 57 (2002) 101–109

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

# On the asymptotic behaviour of unit-root tests in the presence of a Markov trend

Zacharias Psaradakis *

*School of Economics, Mathematics and Statistics, Birkbeck College, University of London, 7-15 Gresse Street, London W1T 1LL, UK*

Received August 2001; received in revised form December 2001

**Abstract**

This paper examines the behaviour of unit-root tests for I(1) time series with drift which is subject to Markov regime changes. It is shown that the asymptotic null distributions of the popular Dickey–Fuller statistics are different from the standard asymptotic distributions obtained under a no-break assumption. Monte Carlo experiments are used to illustrate the finite-sample implications of the theoretical results. 

*Keywords:* Markov chain; Non-stationarity; Structural change; Unit-root test

## 1. Introduction

Recent work on the properties of tests for an autoregressive unit root has cast doubts on the usefulness of many standard tests in situations where the time series of interest is integrated of order one [I(1)] but there is a break in its trend component. Leybourne et al. (1998) and Leybourne and Newbold (2000) have demonstrated that the popular Dickey–Fuller tests (Dickey and Fuller, 1979) have an excessively high Type I error probability when the level or the slope of the trend function is subject to change. Montañés and Reyes (1999), on the other hand, showed that Dickey–Fuller tests (applied to linearly detrended series) are asymptotically correct when there is a break in the level of the trend function.

A limitation of these studies is their maintained assumption that structural change is a one-off event that occurs at a pre-determined date. Such an assumption is at odds with the behaviour of many observed time series, especially economic and financial ones, which often appear to have undergone multiple irregular changes in their structure.

* Tel.: +44-20-7631-6415; fax: +44-20-7631-6416.
*E-mail address:* z.psaradakis@bbk.ac.uk (Z. Psaradakis).


PII: S0167-7152(02)00063-9

The aim of the present paper is to extend the work mentioned above by investigating the asymptotic behaviour of unit-root test statistics in the presence of arbitrarily many stochastic shifts in the trend component. To do so, we take the view that the rate of drift of the time series of interest is subject to discrete shifts that are governed by a finite Markov chain. Such a stochastic specification for the trend was introduced by Hamilton (1989) to characterize the behaviour of non-stationary time series subject to changes in regime, and has attracted a great deal of attention in the econometric and statistical literature, not least because of its considerable success in the modelling of time series with structural changes. For our purposes, in addition to providing a plausible stochastic structure for characterizing situations involving fairly abrupt changes, the Markov trend also has the advantage of negating the need to fix the number and location of possible break-points a priori, allowing as it does for the possibility of arbitrarily many trend breaks at unspecified locations. This is a considerable generalization of the single-break assumption made in earlier work.

Section 2 of the paper contains our main results pertaining to the asymptotic behaviour of unit-root test statistics for time series with a Markov trend. Section 3 illustrates the implications of the theoretical results by means of some Monte Carlo experiments. Section 4 summarizes and concludes.

## 2. Main results

Suppose that $\{X_t,\ t\in\mathbb{Z}^+\}$, $\mathbb{Z}^+:=\{0,1,2,\ldots\}$, is a real-valued stochastic process such that

$$X_t = \eta_t + v_t, \tag{1}$$

$$\eta_t = \eta_{t-1} + \sum_{i=1}^{m} \mu_i I_{\{S_t=i\}}, \tag{2}$$

$$v_t = v_{t-1} + \varepsilon_t, \tag{3}$$

where $\{S_t,\ t\in\mathbb{Z}^+\}$ are random variables taking values in the finite set $\Omega=\{1,2,\ldots,m\}$, $\mu_1,\mu_2,\ldots,\mu_m$ are fixed constants, $\{\varepsilon_t,\ t\in\mathbb{Z}^+\}$ is a white-noise process, the initial values $\eta_{-1}$ and $v_{-1}$ are random and bounded in probability, and $I_{\{A\}}$ is the indicator random variable of the event $A$. The following assumptions about $\{S_t,\ t\in\mathbb{Z}^+\}$ and $\{\varepsilon_t,\ t\in\mathbb{Z}^+\}$ will be maintained throughout the paper.

(A.1) $\{S_t,\ t\in\mathbb{Z}^+\}$ is a stationary, irreducible, aperiodic, temporally homogeneous, first-order Markov chain with state space $\Omega=\{1,2,\ldots,m\}$, invariant distribution $\boldsymbol{\pi}=(\pi_1,\pi_2,\ldots,\pi_m)$, and transition matrix $\mathbf{P}=(p_{ij})_{i,j\in\Omega}$, where $p_{ij}:=\mathsf{P}(S_{t+1}=i|S_t=j)$.

(A.2) $\{\varepsilon_t,\ t\in\mathbb{Z}^+\}$ is a sequence of independent, identically distributed (i.i.d.) random variables such that $\mathsf{E}(\varepsilon_t)=0$, $\mathsf{E}(\varepsilon_t^2)=\sigma^2>0$, and $\mathsf{E}(|\varepsilon_t|^{2+\gamma})<\infty$ for some $\gamma>0$.

(A.3) $\{\varepsilon_t,\ t\in\mathbb{Z}^+\}$ is independent of $\{S_t,\ t\in\mathbb{Z}^+\}$.

The formulation above allows the drift of $\{X_t,\ t\in\mathbb{Z}^+\}$ to be subject to occasional discrete shifts that are governed by a simple Markov chain. This is similar to Hamilton's (1989) Markov model of trend, although we allow for an arbitrary number of states. Notice that, although $\{X_t,\ t\in\mathbb{Z}^+\}$ is non-stationary, we have that $\ddot{X}_t:=X_t-X_{t-1}=\varepsilon_t+\sum_{i=1}^{m}\mu_i I_{\{S_t=i\}}$, so the differenced process $\{\ddot{X}_t,\ t\in\mathbb{N}\}$ is stationary.

Upon observing a finite segment $(X_0, X_1, \ldots, X_n)$ of $\{X_t,\ t \in \mathbb{Z}^+\}$, the Dickey–Fuller tests reject the null hypothesis that $\{X_t,\ t \in \mathbb{Z}^+\}$ is an I(1) process in favour of the alternative of (trend) stationarity when the normalized estimation error $B_n := n(\hat{\phi}_n - 1)$ or the studentized statistic $T_n := (\hat{\phi}_n - 1)/\hat{\omega}_n$ take small values. Here, $\hat{\phi}_n$ is the least-squares estimator of $\phi$ in the autoregressive model

$$X_t = \beta_0 + \beta_1(t - \tfrac{1}{2}n) + \phi X_{t-1} + u_t, \quad t = 1, 2, \ldots, n$$

(where $\{u_t,\ 1 \leqslant t \leqslant n\}$ is assumed to be white noise) and $\hat{\omega}_n$ is the least-squares estimator of the standard deviation of $\hat{\phi}_n$.

Before stating our main result, let us introduce some notations and definitions. As usual, $\mathbf{I}_m$ denotes the identity matrix of order $m$, while $\mathbf{1}_m := (1, 1, \ldots, 1)$ stands for the $m$-element sum vector. Further, we let $\boldsymbol{\mu} := (\mu_1, \mu_2, \ldots, \mu_m)$, $\boldsymbol{\Pi} := \operatorname{diag}(\pi_1, \pi_2, \ldots, \pi_m)$, $\mathbf{P}_\infty := \lim_{k\to\infty} \mathbf{P}^k = \boldsymbol{\pi}^\top \mathbf{1}_m$, and $\mathbf{Q} := \mathbf{P} - \mathbf{P}_\infty$. It is easy to verify that $\mathbf{P}^k - \mathbf{P}_\infty = (\mathbf{P}^{k-1} - \mathbf{P}_\infty)\mathbf{Q} = \mathbf{Q}^{k-1}\mathbf{Q} = \mathbf{Q}^k$ for $k \in \mathbb{N}$.

The following theorem gives the limit distributions of the Dickey–Fuller statistics for processes satisfying the relations in Eqs. (1)–(3).

**Theorem 1.** *Suppose that* $\{X_t,\ 0 \leqslant t \leqslant n\}$ *is a realization of a stochastic process satisfying Eqs.* (1)–(3) *and that assumptions* (A.1)–(A.3) *hold true. Then*

$$B_n \xrightarrow[n\to\infty]{\mathfrak{D}} \left\{ \int_0^1 \mathbb{V}^2(r)\,\mathrm{d}r \right\}^{-1} \left\{ \int_0^1 \mathbb{V}(r)\,\mathrm{d}\mathbb{W}(r) + \delta_1 \right\} \tag{4}$$

*and*

$$T_n \xrightarrow[n\to\infty]{\mathfrak{D}} \left\{ \delta_2 \int_0^1 \mathbb{V}^2(r)\,\mathrm{d}r \right\}^{-1/2} \left\{ \int_0^1 \mathbb{V}(r)\,\mathrm{d}\mathbb{W}(r) + \delta_1 \right\}, \tag{5}$$

*where*

$$\delta_1 := \frac{\boldsymbol{\mu}\mathbf{Q}(\mathbf{I}_m - \mathbf{Q})^{-1}\boldsymbol{\Pi}\boldsymbol{\mu}^\top}{\sigma^2 + \boldsymbol{\mu}[(\mathbf{I}_m - \mathbf{P}_\infty) + 2\mathbf{Q}(\mathbf{I}_m - \mathbf{Q})^{-1}]\boldsymbol{\Pi}\boldsymbol{\mu}^\top},$$

$$\delta_2 := \frac{\sigma^2 + \boldsymbol{\mu}(\mathbf{I}_m - \mathbf{P}_\infty)\boldsymbol{\Pi}\boldsymbol{\mu}^\top}{\sigma^2 + \boldsymbol{\mu}[(\mathbf{I}_m - \mathbf{P}_\infty) + 2\mathbf{Q}(\mathbf{I}_m - \mathbf{Q})^{-1}]\boldsymbol{\Pi}\boldsymbol{\mu}^\top},$$

$$\mathbb{V}(r) := \mathbb{W}(r) - \int_0^1 \mathbb{W}(\tau)\,\mathrm{d}\tau - 12(r - \tfrac{1}{2}) \int_0^1 (\tau - \tfrac{1}{2})\mathbb{W}(\tau)\,\mathrm{d}\tau, \quad r \in [0, 1]$$

*and* $\{\mathbb{W}(r),\ r \in [0, 1]\}$ *is a standard Wiener process.*

**Proof.** Under assumption (A.1), $\{S_t,\ t \in \mathbb{Z}^+\}$ is an ergodic chain with one-dimensional marginal distribution $\boldsymbol{\pi}$ (which satisfies $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}^\top$ and $\mathbf{1}_m\boldsymbol{\pi}^\top = 1$). Hence, putting

$$Y_t := \sum_{i=1}^{m} \mu_i I_{\{S_t = i\}},$$

direct calculations show that

$$\mathsf{E}(Y_t) = \sum_{i=1}^{m} \mu_i \mathsf{P}(S_t = i) = \boldsymbol{\mu}\boldsymbol{\pi}^{\mathsf{T}},$$

$$\mathrm{Var}(Y_t) = \sum_{i=1}^{m} \mu_i^2 \mathsf{P}(S_t = i) - (\boldsymbol{\mu}\boldsymbol{\pi}^{\mathsf{T}})^2 = \boldsymbol{\mu}\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}} - \boldsymbol{\mu}\mathbf{P}_\infty\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}} = \boldsymbol{\mu}(\mathbf{I}_m - \mathbf{P}_\infty)\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}},$$

$$\begin{aligned}\mathrm{Cov}(Y_{t+k}, Y_t) &= \sum_{i=1}^{m}\sum_{j=1}^{m} \mu_i\mu_j \mathsf{P}(S_{t+k} = i | S_t = j)\mathsf{P}(S_t = j) - (\boldsymbol{\mu}\boldsymbol{\pi}^{\mathsf{T}})^2 \\ &= \boldsymbol{\mu}\mathbf{P}^k\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}} - \boldsymbol{\mu}\mathbf{P}_\infty\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}} = \boldsymbol{\mu}\mathbf{Q}^k\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}}, \quad k \in \mathbb{N}.\end{aligned}$$

Now let $\xi_t := Y_t + \varepsilon_t - \boldsymbol{\mu}\boldsymbol{\pi}^{\mathsf{T}}$, so that

$$\ddot{X}_t = \boldsymbol{\mu}\boldsymbol{\pi}^{\mathsf{T}} + \xi_t. \tag{6}$$

Then, recalling that $\{\varepsilon_t,\ t \in \mathbb{Z}^+\}$ is a collection of i.i.d. random variables independent of $\{S_t,\ t \in \mathbb{Z}^+\}$, it is easy to see that $\{\xi_t,\ t \in \mathbb{Z}^+\}$ is a stationary sequence with $\mathsf{E}(\xi_t) = 0$,

$$\mathsf{E}(\xi_t^2) = \sigma^2 + \boldsymbol{\mu}(\mathbf{I}_m - \mathbf{P}_\infty)\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}} \tag{7}$$

and

$$\mathsf{E}(\xi_{t+k}\xi_t) = \boldsymbol{\mu}\mathbf{Q}^k\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}}, \quad k \in \mathbb{N}. \tag{8}$$

Furthermore, we have

$$\begin{aligned}\lim_{n\to\infty} \mathrm{Var}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\xi_t\right) &= \mathsf{E}(\xi_1^2) + 2\sum_{k=1}^{\infty}\mathsf{E}(\xi_{1+k}\xi_1) \\ &= \sigma^2 + \boldsymbol{\mu}(\mathbf{I}_m - \mathbf{P}_\infty)\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}} + 2\sum_{k=1}^{\infty}\boldsymbol{\mu}\mathbf{Q}^k\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}} \\ &= \sigma^2 + \boldsymbol{\mu}[(\mathbf{I}_m - \mathbf{P}_\infty) + 2\mathbf{Q}(\mathbf{I}_m - \mathbf{Q})^{-1}]\boldsymbol{\Pi}\boldsymbol{\mu}^{\mathsf{T}}.\end{aligned} \tag{9}$$

To obtain the last equality in Eq. (9), we have used the fact that the eigenvalues $\lambda_1, \ldots, \lambda_m$ of $\mathbf{P}$ satisfy $\lambda_1 = 1$ and $|\lambda_i| < 1$ for $i \in \{2, \ldots, m\}$ on account of the ergodicity of $\{S_t,\ t \in \mathbb{Z}^+\}$. Consequently, since $\mathbf{Q} = \lim_{k\to\infty}(\mathbf{P} - \mathbf{P}^k)$, the eigenvalues of $\mathbf{Q}$ are zero and $\lim_{k\to\infty}(\lambda_i - \lambda_i^k)$ for $i \in \{2, \ldots, m\}$, and hence $\lim_{k\to\infty}\sum_{j=0}^{k}\mathbf{Q}^j = (\mathbf{I}_m - \mathbf{Q})^{-1}$ exists.

Finally, since $\{S_t,\ t \in \mathbb{Z}^+\}$ is a $\varphi$-mixing sequence with geometrically declining mixing coefficients (cf. Billingsley, 1968, pp. 167–168) and $\{\varepsilon_t,\ t \in \mathbb{Z}^+\}$ is (trivially) $\varphi$-mixing of arbitrarily large size, $\{\xi_t,\ t \in \mathbb{Z}^+\}$ is $\alpha$-mixing of arbitrarily large size. This implies that $\{\xi_t,\ t \in \mathbb{Z}^+\}$ satisfies the mixing condition in Phillips and Perron (1988, p. 336). Thus, by Eqs. (6)–(9) and Theorem 1 of Phillips and Perron (1988) the desired results follow. □

**Remark 1.** It is clear from Theorem 1 that, in the presence of a Markov trend the limit distributions of $B_n$ and $T_n$ are different from the standard Dickey–Fuller asymptotic distributions (which are obtained from Eqs. (4)–(5) with $\delta_1 = 1 - \delta_2 = 0$). The latter are shifted either to the left or to the right relative to the limit distributions in Theorem 1, depending on whether $\mathbf{Q}(\mathbf{I}_m - \mathbf{Q})^{-1}\mathbf{\Pi}$ is positive definite or negative definite, respectively. This implies that the empirical Type I error probability of tests based on $B_n$ and $T_n$ is likely to differ significantly from the nominal level determined on the basis of Dickey–Fuller large-sample theory, yielding misleading conclusions about the order of integration of $\{X_t,\ t \in \mathbb{Z}^+\}$.

**Remark 2.** The asymptotic behaviour of the unit-root statistics $B_n$ and $T_n$ will be unaffected by the presence of the Markov trend if all the sub-dominant eigenvalues of the transition matrix $\mathbf{P}$ are equal to zero. (When $m = 2$, for example, this requires that $p_{22} = 1 - p_{11}$.) In this case, $\mathrm{Cov}(S_{t+k}, S_t) = 0$ for all $k \in \mathbb{N}$ and $\boldsymbol{\mu}\mathbf{Q}(\mathbf{I}_m - \mathbf{Q})^{-1}\mathbf{\Pi}\boldsymbol{\mu}^{\mathsf{T}} = 0$. This is a special yet plausible situation, for there exist time series which appear to be best described as having a trend component which is subject to uncorrelated random shifts (see, e.g., Hansen, 1992).

**Remark 3.** In view of Eq. (6), asymptotically valid unit-root tests can be constructed by relaxing the Dickey–Fuller assumption that $\{\ddot{X}_t,\ t \in \mathbb{N}\}$ is a white-noise process under the null hypothesis. It is not difficult, for example, to show that, since $\{\ddot{X}_t,\ t \in \mathbb{N}\}$ is a stationary $\alpha$-mixing sequence under the conditions of Theorem 1, tests based on statistics like the $Z(\tilde{\alpha})$ and $Z(t_{\tilde{\alpha}})$ statistics of Phillips and Perron (1988) or their modified versions ($MZ_\alpha$ and $MZ_t$) proposed by Perron and Ng (1996) have the correct level asymptotically.

**Remark 4.** In view of the results on the autocovariance structure of $\{\xi_t,\ t \in \mathbb{Z}^+\}$ given in Eqs. (7) and (8), it can be proved, as in Theorem 3 of Zhang and Stine (2001), that $\{\ddot{X}_t,\ t \in \mathbb{N}\}$ admits an ARMA$(p, q)$ representation with $p + 1 \leqslant m$ and $q + 1 \leqslant m$. This implies that a unit-root test based on the so-called augmented Dickey–Fuller (ADF) pseudo-$t$ statistic (see Dickey and Fuller, 1979; Saïd and Dickey, 1984) is likely to have an empirical Type I error probability that is close to the nominal significance level if $n$ is sufficiently large.

## 3. Simulation results

To illustrate the implications of the results in Section 2 numerically, we carried out a few Monte Carlo experiments. In the latter, artificial time series were generated according to Eqs. (1)–(3) with $m = 2$, $\mu_1 = 1$, $\mu_2 \in \{3, 6, 9\}$, $(p_{11}, p_{22}) \in \{(0.8, 0.9), (0.2, 0.3)\}$, and $\varepsilon_t \overset{\mathfrak{D}}{=} \mathcal{N}(0, 1)$.

The finite-sample densities of the unit-root test statistics $B_n$ and $T_n$ when $n = 100$ are shown in Figs. 1 and 2 along with the asymptotic Dickey–Fuller densities. The former were constructed by kernel smoothing of 10,000 Monte Carlo values of the test statistics, using the Gaussian kernel and the bandwidth selector of Sheather and Jones (1991). The asymptotic Dickey–Fuller densities were constructed by similar kernel smoothing of 10,000 Monte Carlo values of $B_n$ and $T_n$, each computed on the basis of 5000 artificial observations from a Gaussian random walk. Consonant with the results in Theorem 1, the finite-sample densities of the test statistics lie to the left of the limit Dickey–Fuller densities when $p_{11} + p_{22} < 1$ , while the converse is true when $p_{11} + p_{22} > 1$. (Note that the eigenvalues of $\mathbf{Q}(\mathbf{I}_m - \mathbf{Q})^{-1}\mathbf{\Pi}$ are zero and $-2\lambda_2(\lambda_2 - p_{11})(\lambda_2 - p_{22})/(\lambda_2 - 1)^3$, where $\lambda_2 = p_{11} + p_{22} - 1$ is the sub-dominant eigenvalue of $\mathbf{P}$; hence, $\delta_1 \leqslant 0$ if $\lambda_2 < 0$ and $\delta_1 \geqslant 0$ if $\lambda_2 > 0$.)

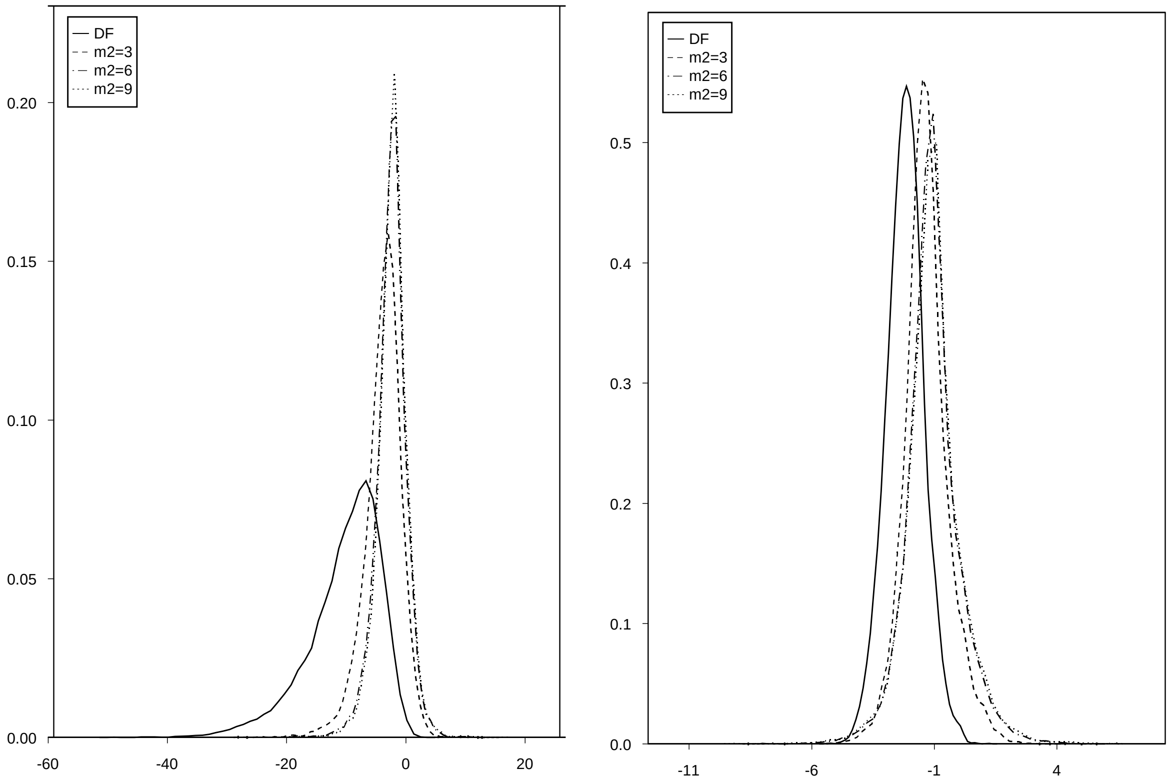


Fig. 1. Asymptotic Dickey–Fuller density of $B_n$ (left) and $T_n$ (right), and exact densities when $n = 100$, $p_{11} = 0.8$, $p_{22} = 0.9$, and $\mu_2 \in \{3, 6, 9\}$.

We also compared the finite-sample Type I error probabilities of tests based on $B_n$ and $T_n$ with those of autocorrelation-robust tests. The latter include tests based on the ADF statistic for an AR$(h+1)$ model with trend (see, e.g., Fuller, 1996, pp. 567–568) and the $MZ_\alpha$ and $MZ_t$ statistics of Perron and Ng (1996). The last two statistics were constructed using an estimate of $\lim_{n\to\infty}\mathrm{Var}(n^{-1/2}\sum_{t=1}^{n}\ddot{X}_t)$ based on an AR$(h+1)$ model (cf. Perron and Ng, 1996, p. 438). For all three tests, $h$ was selected by minimizing the familiar Akaike information criterion (AIC) over the range $0 \leqslant h \leqslant \lfloor 12(n/100)^{1/4}\rfloor$. In addition, we considered tests based on $B_n$ and $T_n$ but with critical values obtained from a sieve bootstrap approximation to the sampling distributions of the test statistics. These bootstrap unit-root tests, proposed by Psaradakis (2001), are based on the idea of approximating the generating mechanism of $\{\ddot{X}_t\}$ by an AR$(h^*)$ model of sufficiently high order, which is then used to resample residuals and generate bootstrap data. In the experiments, the AIC was used to select $h^*$ from the range $1 \leqslant h^* \leqslant \lfloor 12(n/100)^{1/4}\rfloor$.

Table 1 reports the empirical Type I error probability of one-tailed 0.05-level tests when $n = 100$. For tests based on the statistics $B_n$, $T_n$, ADF, $MZ_\alpha$, and $MZ_t$, these were calculated from 10,000 Monte Carlo replications using asymptotic critical values from Fuller (1996, pp. 641–642). For the bootstrap tests, denoted by $B_n^*$ and $T_n^*$, 399 bootstrap iterations for each of 1000 Monte Carlo replications were used to estimate the fifth percentile of the bootstrap distribution of $B_n$ and $T_n$ under

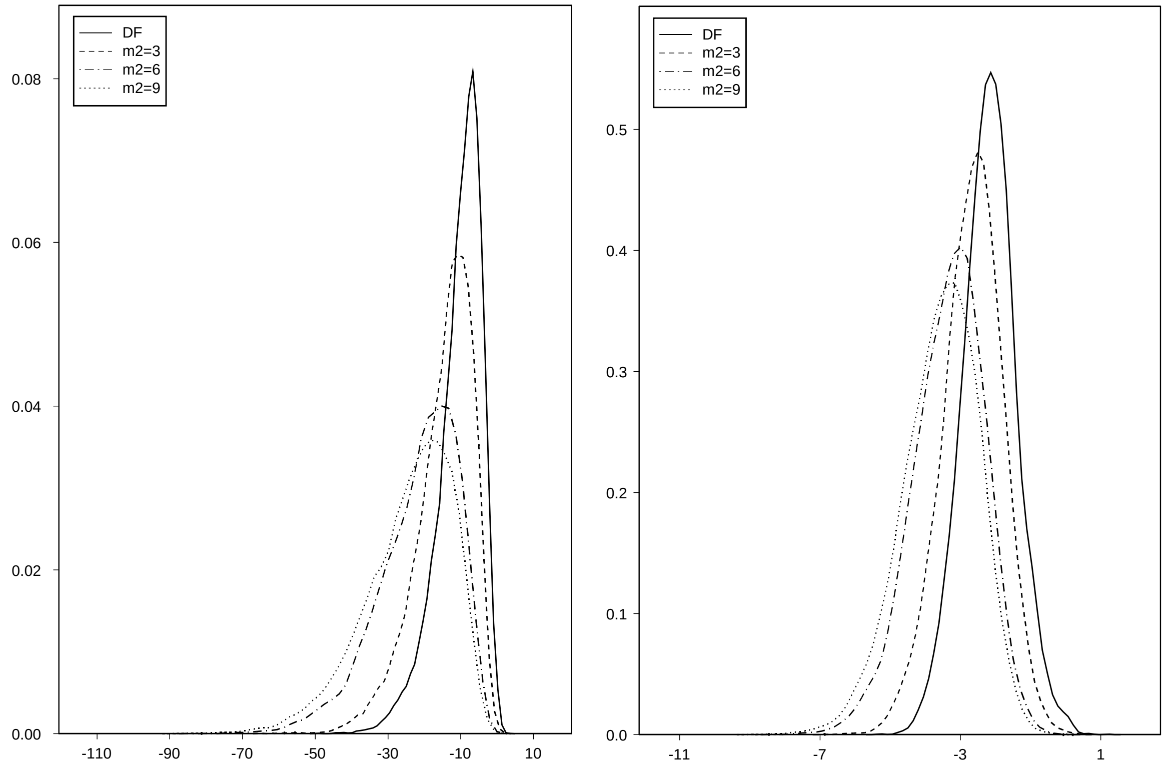


Fig. 2. Asymptotic Dickey–Fuller density of $B_n$ (left) and $T_n$ (right), and exact densities when $n=100$, $p_{11}=0.2$, $p_{22}=0.3$, and $\mu_2 \in \{3, 6, 9\}$.

Table 1
Empirical Type I error probability of 0.05-level tests

| $\mu_2$ | $B_n$ | $T_n$ | ADF | $MZ_\alpha$ | $MZ_t$ | $B_n^*$ | $T_n^*$ |
|---|---|---|---|---|---|---|---|
| $(p_{11}, p_{22}) = (0.8, 0.9)$ | | | | | | | |
| 3 | 0.000 | 0.012 | 0.064 | 0.100 | 0.086 | 0.007 | 0.035 |
| 6 | 0.000 | 0.019 | 0.079 | 0.119 | 0.100 | 0.026 | 0.052 |
| 9 | 0.000 | 0.021 | 0.084 | 0.125 | 0.103 | 0.028 | 0.042 |
| $(p_{11}, p_{22}) = (0.2, 0.3)$ | | | | | | | |
| 3 | 0.165 | 0.182 | 0.102 | 0.083 | 0.064 | 0.022 | 0.020 |
| 6 | 0.421 | 0.439 | 0.079 | 0.063 | 0.050 | 0.018 | 0.016 |
| 9 | 0.512 | 0.524 | 0.074 | 0.055 | 0.043 | 0.013 | 0.011 |

the null hypothesis. As expected, the deviations of the empirical rejection frequencies of the standard Dickey–Fuller tests from the nominal 0.05 value are quite substantial and become more pronounced the larger $\mu_2 - \mu_1$ is. Autocorrelation-robust tests certainly hold an advantage over the Dickey–Fuller tests. (Simple calculations show that the autocovariance structure of $\{\ddot{X}_t\}$ is that of an ARMA(1,1) process with $\mathrm{Var}(\ddot{X}_t)=\sigma^2+(\mu_2-\mu_1)^2\pi_1\pi_2$, $\mathrm{Cov}(\ddot{X}_{t+1},\ddot{X}_t)=(\mu_2-\mu_1)[\pi_1\mu_1(\pi_1-p_{11})-\pi_2\mu_2(\pi_2-p_{22})]$, and $\mathrm{Cov}(\ddot{X}_{t+k},\ddot{X}_t)=(p_{11}+p_{22}-1)^{k-1}\mathrm{Cov}(\ddot{X}_{t+1},\ddot{X}_t)$ for $k \geqslant 2$.) More specifically, tests based on the $MZ_\alpha$ and $MZ_t$ statistics clearly dominate other tests when $(p_{11},p_{22})=(0.2,0.3)$, but they are too liberal when $(p_{11},p_{22})=(0.8,0.9)$. In the latter case, the sieve bootstrap test based on the studentized statistic $T_n$ is the most successful, having empirical Type I error probabilities that are insignificantly different from 0.05 at the 1% significance level. (It is worth mentioning that qualitatively similar results were obtained for $n=200$.)

## 4. Summary

In this paper, we have examined the properties of unit-root tests for I(1) time series which have a Markov trend component. It has been shown that the limit distributions of Dickey–Fuller test statistics are different from the asymptotic distributions given in Dickey and Fuller (1979) and are dependent on nuisance parameters that are associated with the magnitude of the structural changes and the limit properties of the Markov chain that drives these changes. As a consequence, unit-root tests that use as critical values the percentiles tabulated in Fuller (1996, pp. 641–642) are not asymptotically correct in general.

The finite-sample implications of these theoretical results have been illustrated numerically by means of a simulation study. Not surprisingly, Dickey–Fuller tests tend to either over-reject or under-reject the I(1) hypothesis when a Markov trend is present. The problem can be overcome by the use autocorrelation-robust unit-root tests.

## Acknowledgements

I would like to thank an anonymous referee for useful comments.

## References

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Amer. Statist. Assoc. 74, 427–431.

Fuller, W.A., 1996. Introduction to Statistical Time Series, 2nd Edition. Wiley, New York.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrics 57, 357–384.

Hansen, B.E., 1992. The likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP. J. Appl. Econometrics 7, S61–S82.

Leybourne, S.J., Mills, T.C., Newbold, P., 1998. Spurious rejections by Dickey–Fuller tests in the presence of a break under the null. J. Econometrics 87, 191–203.

Leybourne, S.J., Newbold, P., 2000. Behaviour of the standard and symmetric Dickey–Fuller-type tests when there is a break under the null hypothesis. Econometrics J. 3, 1–15.

Montañés, A., Reyes, M., 1999. The asymptotic behaviour of Dickey–Fuller tests under the crash hypothesis. Statist. Probab. Lett. 42, 81–89.

Perron, P., Ng, S., 1996. Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. Rev. Economics Stud. 63, 435–463.

Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335–346.

Psaradakis, Z., 2001. Bootstrap tests for an autoregressive unit root in the presence of weakly dependent errors. J. Time Ser. Anal. 22, 577–594.

Saïd, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71, 599–607.

Sheather, S.J., Jones, M.C., 1991. A reliable data-based bandwidth selection method for kernel density estimation. J. Roy. Statist. Soc. B 53, 683–690.

Zhang, J., Stine, R.A., 2001. Autocovariance structure of Markov regime switching models and model selection. J. Time Ser. Anal. 22, 107–124.